# This Page Is Inserted by IFW Operations and is not a part of the Official Record

# **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

# IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents will not correct images, please do not report the images to the Image Problem Mailbox.

#### -- BRIEF DESCRIPTION OF THE DRAWINGS

This invention will be more fully described with reference to the drawings in which:

1

**Figure 1** depicts the universal code equivalent of the mitochondrial I-Scel gene (SEQ ID NOS:3 AND 4).

**Figure 2** depicts the nucleotide sequence of the invention encoding the enzyme I-Scel and the amino acid sequence of the natural I-Scel enzyme (SEQ ID NOS: 5 and 2).

**Figure 3** depicts the I-Scel recognition sequence and indicates possible base mutations in the recognition site and the effect of such mutations on stringency of recognition (SEQ ID NOS: 6, 7, AND 8).

**Figure 4A-B** is the nucleotide sequence and deduced amino acid sequence of a region of plasmid pSCM525. The nucleotide sequence of the invention encoding the enzyme I-Scel is enclosed in the box (SEQ ID NOS: 9 through 16).

**Figure 5** depicts variations around the amino acid sequence of the enzyme I-Scel (SEQ. ID NO:2).

**Figure 6** shows Group I intron encoding endonucleases and related endonucleases (SEQ ID NOS: 17 through 44).

Figure 7 depicts yeast expression vectors containing the synthetic gene for I-Scel.

Figure 8 depicts the mammalian expression vector PRSV I-Scel.

**Figure 9** is a restriction map of the plasmid pAF100. (See also YEAST, 6:521-534, 1990, which is relied upon and incorporated by reference herein).

**Figures 10A and 10B** show the nucleotide sequence and restriction sites of regions of the plasmid pAF100 (SEQ ID NOS:45 through 50). --

**Figure 11** depicts an insertion vector pTSM $\omega$ , pTKM $\omega$ , and pTT $\omega$  containing the I-Scel site for *E. coli* and other bacteria.

**Figure 12** depicts an insertion vector pTYW6 containing the I-Scel site for yeast.

**Figure 13A-C** depicts an insertion vector PMLV LTR SAPLZ containing the I-Scel site for mammalian cells.

**Figure 14A-B** depicts a set of seven transgenic yeast strains cleaved by I-Scel. Chromosomes from FY1679 (control) and from seven transgenic yeast strains with I-Scel sites inserted at various positions along chromosome XI were treated with I-Scel. DNA was electrophoresed on 1% agarose (SeaKem) gel in 0.25 X TBE buffer at 130 V and 12°C on a Rotaphor apparatus (Biometra) for 70 hrs using 100 sec to 40 sec decreasing pulse times. (A) DNA was stained with ethidium bromide (0.2 μg/ml) and transferred to a Hybond N (Amersham) membrane for hybridization. (B)<sup>32</sup>P labelled cosmid pUKG040 which hybridizes with the shortest fragment of the set was used as a probe. Positions of chromosome XI and shorter chromosomes are indicated.

**Figure 15A-E** depicts the rationale of the nested chromosomal fragmentation strategy for genetic mapping. (A) Positions of *I-Sce*l sites are placed on the map, irrespective of the left/right orientation (shorter fragments are arbitrarily placed on the left). Fragment sizes as measured from PFGE (Fig. 14A) are indicated in kb (note that the sum of the two fragment sizes varies slightly due to the limit of precision of each measurement). (B) Hybridization with the probe that hybridizes the shortest fragment of the set determines the orientation of each fragment (see Fig. 14B). Fragments that hybridize with the probe (full lines) have been placed arbitrarily to the left. (C) Transgenic yeast strains have been ordered with increasing sizes of hybridizing chromosome fragments. (D) Deduced *I-Sce*l map with minimal and maximal size of intervals indicated in kb (variations in some intervals are due to limitations of PFGE measurements). (E) Chromosome subfragments are used as probes to assign each cosmid clone to a given map interval or across a given *I-Sce*l site.

**Figure 16A-C** depicts mapping of the *I-Sce*l sites of transgenic yeast strains by hybridization with left end and right end probes of chromosome XI. Chromosomes from FY1679 (control) and the seven transgenic yeast strains were digested with *I-Sce*l. Transgenic strains were placed in order as explained in Fig. 15. Electrophoresis conditions were as in Fig. 14. <sup>32</sup>P labelled cosmids pUKG040 and pUKG066 were used as left end and right end probes, respectively.

Figure 17A-H depicts mapping of a cosmid collection using the nested chromosomal fragments as probes. Cosmid DNAs were digested with EcoRI and electrophoresed on 0.9% agarose (SeaKem) gel at 1.5 V/cm for 14 hrs, stained with ethidium bromide and transferred to a Hybond N membrane. Cosmids were placed in order from previous hybridizations to help visualize the strategy. Hybridizations were carried out serially on three identical membranes using left end nested chromosome fragments purified on PFGE (see Fig. 16) as probes. A: ethidium bromide staining (ladder is the BRL "1kb ladder"), B: membrane #1, probe: Left tel to A302 site,

C: membrane #1, probe: Left tel to M57 site, D: membrane #2, probe: Left tel to H81 site, E: membrane #2, probe: Left tel to T62 site, F: membrane #3, probe: Left tel to G41 site, G: membrane #3, probe: Left tel to D304 site, H: membrane #3, probe: entire chromosome XI.

**Figure 18** depicts a map of the yeast chromosome XI as determined from the nested chromosomal fragmentation strategy. The chromosome is divided into eight intervals (with sizes indicated in kb, see Fig. 15D) separated by seven *I-Sce*I sites (E40, A302 ...). Cosmid clones falling either within intervals or across a given *I-Sce*I site are listed below intervals or below interval boundaries, respectively. Cosmid clones that hybridize with selected genes used as probes are indicated by letters (a-i). They localize the gene with respect to the *I-Sce*I map and allow comparison with the genetic map (top).

**Figure 19A-B** depicts diagrams of successful site directed homologous recombination experiments performed in yeast.

Figure 20A-E. Experimental design for the detection of HR homologous recombination (HR) induced by I-Sce I. a) Maps of the 7.5 kb tk -PhleoLacZ retrovirus (G-MtkPL) and of the 6.0 kb PhleoLacZ retrovirus (G-MPL), SA is splice acceptor site. G-MtkPL sequences (from G-MtkPL virus) contains PhleoLacZ fusion gene for positive selection of infected cells (in phleomycin-containing medium) and tk gene for negative selection (in gancyclovir-containing medium). G-MPL sequences (from G-MPL virus) contains only PhleoLacZ sequences. b) Maps of proviral structures following retroviral integration of G-MtkPL and G-MPL. I-Sce I PhleoLacZ LTR duplicates, placing I-Sce I PhleoLacZ sequences in the 5' LTR. The virus vector (which functions as a promoter trap) is transcribed (arrow) by a flanking cellular promoter, P. c) I-Sce I creates two double strand breaks (DSBs) in host DNA liberating the central segment and leaving broken chromosome ends that can pair with the donor plasmid, pVRneo (d). e) Expected recombinant locus following HR.

**Figure 21**<u>A-B.</u> **A.** Scheme of pG-MPL. SD and SA are splice donor and splice acceptor sites. The structure of the unspliced 5.8 kb (genomic) and spliced 4;2 kb transcripts is shown below. Heavy bar is  $^{32}$ P radiolabelled LacZ probe (**P**). B. RNA Northern blot analysis of a pG MLP transformed  $\psi$ -2 producer clone using polyadenylated RNA. Note that the genomic and the spliced mRNA are produced at the same high level.

**Figure 22A-B. A.** Introduction of duplicated I-Sce I recognition sites into the genome of mammalian cells by retrovirus integration. Scheme of G-MPL and G-MtkPL proviruses which illustrates positions of the two LTRs and pertinent restriction sites. The size of *Bcl* I fragments and of I-Sce I fragments are indicated. Heavy bar is <sup>32</sup>P radiolabelled *LacZ* probe (**P**). **B.** Southern blot analysis of cellular DNA from NIH3T3

fibroblasts cells infected by G-MtkPL and PCC7-S multipotent cells infected by G-MPL. Bcl I digests demonstrating LTR mediated PhleoLacZ duplication; I-Sce I digests demonstrating faithful duplication of I-Sce I sites.

Figure 23A-B. Verification of recombination by Southern. A.: Expected fragment sizes in kilobase pairs (kb) of provirus at the recombinant locus. 1) the parental proviral locus. Heavy bar (P) is <sup>32</sup>P radioactively labelled probe used for hybridization. 2) a recombinant derived after cleavage at the two I-Sce I sites followed by gap repair using pVR neo (double-site homologous recombination, DsHR). 3) a recombination event initiated by the cleavage at the I-Sce I sites in the left LTR (single-site homologous recombination, SsHR). B.: Southern analysis of DNA from NIH3T3/G-MtkPL clones 1 and 2, PCC7-S/G-MPL clones 3 and 4 and transformants derived from cotransfection with pCMV(I-Sce I+) and pVRneo (1a, 1b, 2a, 3a, 3b and 4a). *Kpn* I digestion of the parental DNA generates a 4.2 kb fragment containing *LacZ* fragment. Recombinants 1a and 3a are examples of DsHR Recombinants 1b, 2a, 3b and 4a are examples of SsHR.

Figure 24<u>A-B</u>. Verification of recombination by Northern blot analyses. A.: Expected structure and sizes (in kb) of RNA from PCC7-S/G-MPL clone 3 cells before (top) and after (bottom) I-Sce I induced HR with pVRneo.I Heavy bars P1 and P2 are <sup>32</sup>P radioactively labelled probes. B.: Northern blot analysis of the PCC7-S/G-MPL clone 3 recombinant (total RNA). Lane 3 is parental cells, lane 3a recombinant cells. Two first lanes were probed with *LacZ* P1, two last lanes are probed with *neo* P2. parental PCC7-S/G-MPL clone 3 cells express a 7.0 kb *LacZ* RNA as expected of trapping of a cellular promoter leading to expression of a cellular-viral fusion RNA. The recombinant clone does not express this *Lacz* RNA but expresses a *neo* RNA of 5.0 kb, corresponding to the size expected for an accurate replacement of *PhleoLacZ* by *neo* gene.

Figure 25A-C. Types of recombination events induced by I-Sce I DSBs, a) Schematic drawing of the structure of the recombination substrate. The G-MtkPL has provirus two LTRs, each containing an I-Sce I recognition site and a *PhleoLacZ* gene. The LTRs are separated by viral sequences containing the *tk* gene. The phenotype of G-MtkPL containing cells is Phleo<sup>R</sup>, GIs<sup>S</sup>, β-Gal±. b) Possible modes of intra-chromosomal recombination. 1) The I-Sce I endonuclease cuts the I-Sce I site in the 5'LTR. The 5' part of U3 of the 5'LTR can pair and recombine with it homologous sequence in the 3'LTR (by single-strand annealing (SSA). 2) The I-Sce I endonuclease cuts the I-Sce I site in the 3'LTR. The 3' part of U3 of the 3'LTR can pair and recombine with its homologous sequence in the 5'LTR (by SSA). 3) The I-Sce I endonuclease cuts I-Sce I sites in the two LTRs. The two free ends can relegate (by an end-joining mechanism). The resulting recombination product in each of the three models is a solitary LTR (see right side). No modification would occur in the cellular sequences flanking the integration site. c) The I-Sce I endonuclease cuts the I-Sce I sites in the

two LTRs. The two free ends can be repaired (by a gap repair mechanism) using the homologous chromosome. On the right, the resulting recombination product is the deletion of the proviral integration locus.

Figure 26A-C. Southern blot analysis of DNA from NIH3T3/G-MtkPL 1 and 2, and *PhleoLacZ* recombinants derived from transfections with pCMV(I-Sce I+) selected in Gancyclovir containing medium. a) Expected fragment sizes in kilobase pair (kbp) of parental provirus after digestion with *Pst* I endonuclease. *Pst* I digestion of the parental DNA NH3T3/G-MtkPL 1 generates two fragments of 10 kbp and of the parental NIH3T3/G-MtkPL 2 two fragments of 7 kbp and 9 kbp. b) Southern blot analysis of DNA digested by *Pst* I from NIH3T3/G-MtkPL 1, and recombinants derived from transfection with pCMV(I-Sce I+) (1.1 to 1.5). c) Southern blot analysis of DNA digested by *Pst* I from NIH3T3/G-MtkPL 2, and recombinants derived from transfection with pCMV(I-Sce I+) (2.1 to 2.6). Heavy bar is <sup>32</sup>P radiolabelled *LacZ* probe (P).

Figure 27A-C. Southern blot analysis of DNA from NIH3T3/G-MtkPL 1 and 2, and *PhleoLacZ*<sup>†</sup> recombinants derived from transfections with pCMV(I-Sce I+) and pCMV(I-Sce I-) and selection in Phleomycin and Gancyclovir containing medium.

a1) Expected fragment sizes in kbp of parental provirus after digestion with *Pst* I or *Bcl* I endonuclease. *Pst* I digestion of the parental DNA NIH3T3/G-MtkPL 1 generates two fragments of 10 kbp. *Bcl* I digestion of the parental DNA NIH3T3/G-MtkPL 2 generates three fragments of 9.2 kbp, 7.2 kbp and 6.0 kbp. a2) Expected fragment sizes in kbp of recombinants after digestion with *Pst* I or *Bcl* I endonuclease. *Pst* I digestion of DNA of the recombinant derived from NIH3T3/G-MtkPL 1 generates one fragment of 13.6 kbp. *Bcl* I digestion of the DNA of the recombinants derived from NIH3T3/G-MtkPL 2 generates two fragments of 9.2 kbp and 6.0 kbp. b) Southern blot analysis of DNA from NIH3T3/G-MtkPL 1, and recombinants derived from transfection with pCMV(I-Sce I-) and pCMV(I-Sce I+) (1c, 1d). c) Southern analysis of DNA from NIH3T3/G-MtkPL 2, and transformants derived from transfection with pCMV(I-Sce I-) (2a, 2b) and pCMV(I-Sce I+) (2c to 2h). Heavy bar is <sup>32</sup>P radiolabelled *LacZ* probe (P).

- **Figure 28.** Figure 28 is a diagram illustrating the loss of heterozygosity by the insertion or presence of an I-Sce I site, expression of the enzyme I-Sce I, cleavage at the site, and repair of the double strand break at the site with the corresponding chromatid.
- **Figure 29.** Figure 29 is a diagram illustrating conditional activation of a gene. An I-Sce I site is integrated between tandem repeats, and the enzyme I-Sce I is expressed. The enzyme cleaves the double stranded DNA at the I-Sce I site. The double strand break is repaired by single stand annealing, yielding an active gene.
- **Figure 30.** Figure 30 is a diagram illustrating one step rearrangement of a gene by integration of an I-Sce I site or by use of an I-Sce I site present in the gene. A

plasmid having either one I-Sce I site within an inactive gene, or two I-Sce I sites at either end of an active gene without a promoter, is introduced into the cell. The cell contains an inactive form of the corresponding gene. The enzyme I-Sce I cuts the plasmid at the I-Sce I sites, and recombination between the chromosome and the plasmid yields an active gene replacing the inactive gene.

- **Figure 31.** Figure 31 is a diagram illustrating the duplication of a locus. An I-Sce I site and a distal part of the locus are inserted into the gene by classical gene replacement. The I-Sce I site is cleaved by I-Sce I enzyme, and the break is repaired by homologous sequences. This results in duplication of the entire locus.
- **Figure 32.** Figure 30 32 is a diagram illustrating the deletion of a locus. Two I-Sce I sites are added to flank the locus to be deleted. The I-Sce I enzyme is expressed, and the sites are cleaved. The two remaining ends recombine, deleting the locus between the two I-Sce I sites.
- **Figure 33.** Figure 33 is a diagram of plasmid pG-MtkΔPAPL showing the restriction sites. The plasmid is constructed by deletion of the polyadenylation region of the tk gene from the pGMtkPL plasmid. --

On page 18, replace the last paragraph with the following amended paragraph:

-- The enzyme I-Scel has a known recognition site. (ref. 14.) The recognition site of I-Scel is a non-symmetrical sequence that extends over 18 bp as determined by systematic mutational analysis. The sequence reads: (arrows indicate cuts)

On pages 47-48, replace the paragraph bridging these pages with the following amended paragraph:

-- -e- The supernatant of this clone was used to infect other mouse cells (1009) by spreading 10<sup>5</sup> virus particles on 10<sup>5</sup> cells in DMEM medium with 10% fetal calf serum and 5 mg/ml of "polybrain" polybrene (hexadimethrine bromide). Medium was replaced 6 hours after infection by the same fresh medium. --

On pages 71-72, replace the paragraph bridging these pages with the following amended paragraph:

-- This example describes the use of the I-Sce I meganuclease (involved in intron homing of mitochondria of the yeast Saccharomyces cerevisiae) (6B, 28B) to induce DSB and mediate recombination in mammalian cells. I-Sce I is a very rare-cutting restriction endonuclease, with an 18 bp long recognition site (29B, 22B). In vivo, I-Sce I endonuclease can induce recombination in a modified yeast nucleus by initiating a specific DBS leading to gap repair by the cell (30B, 17B, 21B). Therefore, this approach can potentially be used as a means of introducing specific DSB in chromosomal target DNA with a view to manipulate chromosomes in living cells. The I-Sce I-mediated recombination is superior to recombinase system [11B] for chromosome engineering since the latter requires the presence of target sites on both host and donor DNA molecules, leading to reaction that is reversible. --

On pages 72-73, replace the paragraph bridging these pages with the following amended paragraph:

-- pG-MtkPL was obtained in five steps: (I) insertion of the 0.3 kbp *Bgl II-Sma* I fragment (treated with Klenow enzyme) of the Moloney Murine Leukemia Virus (MoMuLV) *env* gene (25B) containing a splice acceptor (SA) between the *Nhe* I and *Xba* I sites (treated with Klenow enzyme), in the U3 sequence of the 3'LTR of MoMuLV, in an intermediate plasmid. (II) Insertion in this modified LTR of a 3.5 kbp *Nco* I-*Xho* I fragment containing the *PhleoLacZ* fusion gene [13B] (from pUT65; Cayla Laboratory, Zone Commerciale du Gros, Toulouse, France) at the *Xba* I site next to SA. (III) Insertion of this 3'LTR (containing SA and *PhleoLacZ*), recovered by *Sal* I-*EcoR* I double digestion in the p5'LTR plasmid (a plasmid containing the 5'LTR up to the nucleotide no 563 of MoMuLV [12B]) between the *Xho* I and the *EcoR* I site. (IV) Insertion of a *synthetic* I-*Sce* I recognition site into the *Nco* I site in the 3'LTR (between SA and *PhleoLacZ*), and (V) insertion (antisense to the retroviral genome) of the 1.6 kbp *tk* gene with its promoter with linker adaptators at the *Pst* I site of pG-MPL. --

On page 80, replace the paragraph beginning on line 5 with the following amended paragraph:

-- The generation of *tk-lPhleoLacZ* cells is probably a consequence of either a homo-allelic and/or an ectopic gene conversion event (36B). Isolation and detailed molecular analysis of the proviral integration sites will provide information on the relative frequency of each of these events for the resolution of chromosomal DSBs by the cell. This quantitative information is important as, in mammalian cells, the high redundancy

of genomic sequences raises the possibility of a repair of DSBs by ectopic homologous sequences. Ectopic recombination for repair of DSBs may be involved in genome shaping and diversity in evolution [29]. --

After making the above amendments, please move pages 50-58 and insert these pages after page 88 and before page 89 of the application.

Renumber pages 50-58 as pages 80-88, respectively.

Renumber pages 59-88 as pages 50-79, respectively.

On pages 89-93, replace the section entitled "References" with the following amended section.

#### -- References

- 1 B. Bernstein, N., Pennell, N., Ottaway, C. A. and Shulman, M.J.
- 1992. Gene replacement with one-sided homologous recombination. Mol. Cell Biol. 12: 360-367.
- 2 <u>B</u>. Bonnerot, C., Legouy, E., Choulika, A. and Nicolas, J.-F. 1992. Capture of a cellular transcriptional unit by retrovirus: mode of provirus activation in embryonal carcinoma cells. J. Virol. 66: 4982-4991.
- 3 <u>B</u>. Bonnerot, C., and Nicolas, J.-F. 1993. Application of LacZ gene fusions to post-implantation development. In "Methods in Enzymology: Guide to techniques in mouse development". 451-469. Wassarman, P.M., DePamphilis, M.L.
- 4 <u>B</u>. Brenner, D.A., Smogocki, A. and Camerini-Otero, R.D. 1986. Double-strand gap repair results in homologous recombination in mouse L cells. Proc. Natl. Acad. Sci. USA. 83: 1762-1766.
- 5 <u>B</u>. Capecchi, M.R. 1989. Altering the genome by homologous recombination. Science. 244: 1288-1292.
- 6 <u>B</u>. Jacquier, A. and Dujon, B. 1985. An intron encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. Cell. 41: 383-394.

- 7 <u>B</u>. Jakob, H. and Nicolas, J.F. 1987. Mouse Tertocarcinoma Cells. In "Methods in Enzymology: Cell lines for genetic analysis". 66-81. Gottesman, M.H., ed., Academic Press.
- 8 <u>B</u>. Jessberger, R. and Berg, P. 1991. Repair of deletions and double-strand gaps by homologous recombination in a mammalian in vitro system. Mol. Cell Biol. 11: 445-457.
- 9 <u>B</u>. Kilby, N.J., Snaith, M.R. and Murray, J.A.H. 1993. Site-specific recombinases: tools for genome engineering. Reviews. 9: 413-421.
- 10 <u>B</u>. Lin, F.L.M., Sperle, K. and Sternberg N. 1990. Repair of double-stranded DNA breaks by homologous DNA fragments during transfer of DNA into mouse L cells. Mol. Cell Biol. 10:113-119.
- 11 <u>B</u>. Lin, F.L.M., Sperle, K. and Sternberg N. 1990. Intermolecular recombination between DNAs introduced into mouse L cells is mediated by a nonconservative pathway that leads to crossover products. Mol. Cell Biol. 10: 103-112.
- 12 <u>B</u>. Lin, F.L.M., Sperle, K. and Sternberg N. 1990. Intermolecular recombination between DNAs introduced into mouse L cells is mediated by a nonconservative pathway that leads to crossover products. Mol. Cell. Biol. 10: 103-112.
- 13 <u>B</u>. Mann, R., Mulligan, R.C. and Baltimore, D. 1983. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. Cell. 33: 153-160.
- 14 <u>B</u>. Mansour, S.L., Thomas K.R. and Capecchi, M.R. 1988. Disruption of the proto-oncogène int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable gene. Nature. 336:348-352.
- 15 <u>B</u>. Mulsant, P., Gatignol, A., Dolens, M. and Tiraby, G. 1988. Somat. Cell. Mol. Genet. 14: 243-252.
- 16 <u>B</u>. Nicolas, J. F. and Rubenstein, J. 1987. Retroviral vectors. Boston London Durban Singapore Sydney Toronto Wellington, Butterworths.

- 17 <u>B</u>. Plessis, A., Perrin, A., Haber, J.E. and Dujon, B. 1992. Site specific recombination determined by I-Sce I, a mitochondrial group I intron-encoded endonuclease expressed in the yeast nucleus. Genetics 130:451-460.
- 18 <u>B</u>. Sauer, B. and Henderson, N. 1988. Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. Prac. Natl. Acad. Sci. USA. 85: 5166-5170.
- 19 <u>B</u>. Seidman, M.M. 1987. Intermolecular homologous recombination between transfected sequences in mammalian cells is primarily nonconservative. Mol. Cell. Biol. 7: 3561-3565.
- 20 <u>B</u>. Smithies, O., Gregg, R.G., Boggs, S.S., Koralewski, M.A. and Kucherlapati, R.S. 1985. Insertion of DNA sequences into the human chromosomal B-globin locus by homologous recombination. Nature. 317: 230-234.
- 21 <u>B</u>. Szostak, J. W., Orr-Weaver, T.L. and Rothstein, R.J. 1983. The double-strand break repair model for recombination. Cell. 33: 25-35.
- 22 <u>B</u>. Thierry, A., Perrin, A., Boyer, J., Fairhead, C., Dujon, B., Frey, B. and Schmitz, G. 1991. Cleavage of yeast and bacteriophage T7 genomes at a single site using the rare cutter endonuclease I-Sce I. Nucleic Acids Res. 19: 189-90.
- 23 <u>B</u>. Tybulewicz, V.L. J., Crawford, C.E., Jackson, P.K., Bronson, R.T. and Mulligan, R.C. 1991. Neonatal Lethality and Lymphopenia in Mice with a Homozygous Disruption of the c-abl Proto-Oncogene. Cell 65: 1153-1163.
- 24 <u>B</u>. Varmus, H. and Brown, P. 1989. Retroviruses <u>in Mobile DNA, 58-108 (Douglas E. Berg and Martha H. Howe eds).</u>
- 25 <u>B</u>. Weiss, R., Teich, N., Varmus, H. and Coffin, J. 1985. RNA tumor viruses. Molecular Biology of tumor viruses. Second Edition. 2) Supplements and appendixes. Cold Spring Harbor Laboratory. 1-1222.
- 26 <u>B</u>. Weiss, R., Teich, N., Varmus, H. and Coffin, J. 1985. RNA tumor viruses. Molecular Biology of tumor viruses. Second Edition. 2) Supplements and appendixes. Cold Spring Harbor Laboratory. 1-1222.

- 27 <u>B</u>. Phillips J. and Morgan W. 1994. Illegitimate recombination induced by DNA double-strand breaks in mammalian chromosomes. Molecular and Cellular Biology 14:5794-5803.
- 28 <u>B</u>. Dujon B. 1989. Group I introns are mobile genetic elements: facts and mechanistic speculations-a review. Gene 82:91-114.
- 29 <u>B</u>. Colleaux L., D'Aurio L., Galibert F. and Dujon B. 1988. Recognition and cleavage site of the intron-encoded omega transposase. Proc Natl Acad Sci USA 85:6022-6.
- 30 <u>B</u>. Fairchild C. and Dujon B. Consequences of unique double-stranded breaks in yeast chromosomes: death or homozygosis. Molecular general genetics 240:170-180.
- 31 <u>B</u>. Pfeiffer P., Thode S., Hancke J. and Vielmetter W. 1994. Mechanism of overlap information in nonhomologous DNA end joining. Molecular and Cellular Biology 14:888-895.
- 32 <u>B</u>. Mezard C. and Nicholas A. 1994. Homologous, homeologous, and illegitimate repair of double-strand breaks during transformation of a wild-type strain and a rad52 Mutant strain of Saccharomyces cerevisiae. Molecular and Cellular Biology 14:1278-1292.
- 33 <u>B</u>. Feaver W.J., Svejstrup J.Q., Bradwell L., Bradwell A.J., Buratowski S., Gulyas K., Donahue T.F., Friedberg E.C. and Kornberg R.D. 1993. Dual Roles of a Multiprotein Complex from S. cerevisiae in transcription and DNA Repair. Cell 75:1379-1387.
- 34 <u>B</u>. Kramer K., Brock J., Bloom K., Moore K. and Haber J. 1994. Two different types of double-strand breaks in Saccharomyces ceerevisiae are repaired by similar RAD52 independent, nonhomolgous recombination events. Molecular and Cellular Biology 14:1293-1301.
- 35 <u>B</u>. Weiffenbach B. and Haber J. 1981. Homothallelic mating type switching generates lethal chromosomes breaks in rad52 strains of Saccharomyces cerevisiae. Molecular and Cellular Biology 1:522-534.

- 36 <u>B</u>. Nassif N., Penney J., Pal S., Engels W. and Gloor G. 1994. Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. Molecular and cellular biology 14:1643-1625.
- 37 <u>B</u>. Charlesworth B., Sniegowski P. and Stephan W. 1994. The evolutionary dynamics of repetitive DNA in eucaryotes. Nature 371:215-220. --

After page 93, and before page 94, insert the attached pages titled "SEQUENCE LISTING".

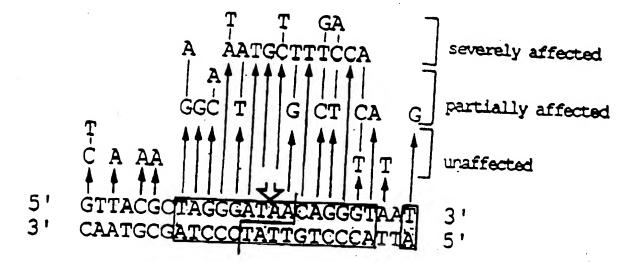
The Universal Code Equivalent of the Mitochondrial 1-Sce I Gene.

# The synthetic I-Sce I gene

Born 41

CCGGATCCATE CAT ATG AAA AAC ATC AAA AAA AAC CAG GTA ATG AAC CTG GST CCG AAC TCT AAA CTG CTG AAA GAA TAC AAA TCC CAG CTG ATC GAA CTG AAC ATC GAA CAG TTC GAA GCA GGT ATC GGT CTG ATC CTG GGT GAT GGT TAC ATC CGT TCT CGT GAT GAA GET AAA ACC TAC 1. TGT ATG CAG TTC GAG TER AAA AAC AAA SCA TAC ATG GAC CAC GTA TGT CTG CTG TAC GAT '**K** N K A Y CAG TEG GTA CTO TEC CEG CEE CAC AMA AMA CAM COT GIT AME CAC CTO GOT AME CTO GTA ATC ACC THE BOC GCC CAR ACT TTC AAA CAC CAA BCT TTC AAC AAA CTE GCT AAC CTG TTC ATC GTT AAC AAC AAA AAA ACC ATC CCB AAC AAC CTB GIT GAA AAC TAC CTB ACC CCG AIG TET CTG GCA TAC TGG TIC ATG GAT GAT GGT AAA TGG GAT TAC AAC AAA AAC TET ACC П AAC AAA TCG ATC GTA CTG AAC ACC CAG TCT TTC ACT TTC GAA GAA GTA GAA TAC CTG GTT ANG GET CTG CGT AND AMA TTC CAM CTG AMC TGT TAC GTA AMA ATC MAC AMA CCG ATC ATC TAC ATC BAT TOT ATG TOT TAC CTB ATC TAC AAC CTE ATC AAA CCB TAC CTE ۲ ATC CCE CAG ATE ATE TAC AAA CTG CCE AAC ACT ATC TCC TCU GAA ACT TIC CTB AAA TAA TAAGTCGACTGCAGGATCCGGTAAGTAA Sail Patl Bank!

1 and 2: These amino acids are absolutely necessary to produce catalytic activity. Other substitutions are possible, such as deletions of the 10 first amino acids.



1447 CONSIGNATIA PROMOTE ACCORDINATION CONCENSATION AND COMPANION THE ARTHUR CONCENTION CONCENTRATION CONCENTRATIO					1																	
1419 THARMSTEPHINGENSTITISCENSCENTARY AND CORE OF THE ATM CORE	166		عىع	1	1000	2571	<b>1370</b>	عدت	œ	can	عمد			~.~								
1819 OFT BYS GAA ANA COC CAG-CAA COC SOC CITY THE ACT OFT COT GOC CITY TO COC GITY TO COC ACT ACT ACT COC ACT ACT ACT COC ACT ACT ACT ACT ACT ACT ACT ACT ACT AC	1245	-	***	-~	<u> </u>											٥	- Maid	بمتحد	MC S	æ <del></del> .	::x:c:	11144
1819 OFT BYS GAA ANA COC CAG-CAA COC SOC CITY THE ACT OFT COT GOC CITY TO COC GITY TO COC ACT ACT ACT COC ACT ACT ACT COC ACT ACT ACT ACT ACT ACT ACT ACT ACT AC				CCI	1100	SC TT	rcæ:	2162	स्य स्त	بنناة	3	czce	7111	1361	E AS	• •	= 510	: Age	3 364	. عد د		
1879 THE TEA CAT GTT CTT TOT TOC GTT ATC CTC TOA INTOTOGRATAACCGTA THE CATCHTCATCHCTCACC 134 12															-	•	v		•		•	
1879 THE TEA CAT GTT CTT TOT TOC GTT ATC CTC TOA INTOTOGRATAACCGTA THE CATCHTCATCHCTCACC 134 12	101	œ.	, y24	, an	r m	r cec	CA	5-CA	<b>A</b> : <b>CC</b>	: 22	: =	7 77	1 10			: œ		· -				- 0
1879 FAU TEA CAT GIT CIT TOT TOE OF ATC CIT ATC CIT TAX CIT TAX CAT GIT TOT TOE TOE OF ATC CIT TAX CIT	•	, ,	-	E	K	R	Q	Q	R	3	L	F	7	V	•	4			;	a arci	: :==	
1948 THU USE TOO CAT OUT CAT OUT CAN ACT TO TOO TOO SEE BE BE TO BE ACT OF TO THE ACT TO	:875	74	: 12	CAS	<u> </u>	2 021	•			- 1~	,ـــ						_		_	••	•	•
1916 THU USE TOO CAT OUT CAT OUT OAN ATT TO TOO TOU GOT AND THE ORD ATT TO TOO ATT TO TOO AND ATT TO TOO TOO AND ATT TO TOO TOO ATT TO TOO ATT TOO ATT TO TOO ATT TOO ATT TO TOO ATT TO TOO ATT TOO ATT TO TOO ATT TO TOO ATT TO TOO AT	20	C		*	y	L	• •	: : -	v	Ī	. C.\		117	. rest		ucc	- 1	, 700	· (	4	7546	1347
2016 COCCTOTOCOCOSCOCITICACTUATIA ATC CAG CTO SCA CCA CAG GTT TOC TOA CTO TAA ACC TO TAA ACC TO TAA ACC TO TAA ACC TO TAA ACC TAC TAC TO TAA ACC TAC TAC TAC TAC TAC TAC TAC TAC										_	-					•						
2016 COCCTOTOCOCOSCOCITICACTUATIA ATC CAG CTO SCA CCA CAG GTT TOC TOA CTO TAA ACC TO TAA ACC TO TAA ACC TO TAA ACC TO TAA ACC TAC TAC TO TAA ACC TAC TAC TAC TAC TAC TAC TAC TAC	1344			W-35	. W.C.	E A CE	:	rcey(	CLX	X	(CC)	<b>GTO</b>	VCIO	ديع	æ	e ca	uca	æ	::::	-1	-	222
2984 COC CAG TWA SCHMANGCANTIA ATG ION OTTAGTCACCTATAGGCACCTTACACCTT ATG 113 C	2020	coc	ETCI	œ	:0050	EIR	2000	AET	ATT (	ATC	-	. ~										•
2008 COC CAG TON GENERALSECANTIA ATG IOM STEMETICACTUATIAGECATCRAGGETTIACACTTI ATG 110 G G C C C C C C C C C C C C C C C C C	1								,	'n	, <u>J</u>	L	, <u>~</u>	l CW	CAG	GT	120	:::				
2144 ACT OF COC CAT CAT ATT GAT AND AND TOT SON ANT GTO AND AND TOT AND CALT TYPO AND AND AND AND CAT THE GAT CAT GTO AND AND AND AND CAT THE GAT CAT GTO AND AND AND AND CAT THE GAT CAT GTO AND AND CAT THE GAT CAT CAT CAT CAT CAT CAT CAT CAT CAT C	2000	-	~	-						_	-				. •		-	•		_		12
2114 AC CTC CCC CCC CCC ATC CCC CAD CAD CTC CCC CCC CCC CCC CCC CCC CCC CCC CC	13	•			- GC (			ATTA	ATG	IGA	GITA	محتو		7117	OCC )	$\alpha$	AGGC	TTTA	CACT	71 A	:5	2164
2229 MCC ATC ATC ATC ATC ATC CAT CAT CAT TAX CAMCTEATCHATCHATCHATCHATCHATCHATCHATCHATCHATCH									-	•												:
2229 MCC ATC ATC ATC ATC ATC CAT CAT CAT TAX CAMCTEATCHATCHATCHATCHATCHATCHATCHATCHATCHATCH	2145	•	, CCC	œ	ces	ASC	: :::	151	C DOL	AIT	314	AGC	- CLA	TAA	مت	TITE	AC AC	AGEA	MC)	ce	ATG	2324
2296 AGTTANATIGCTAACGCACTCAGCACTCITICT ATO ANA TOT AND ANT GOOD CTC ATC CTC ACC 2216.2  23164 ACC STC ACC CTG GAT GET GTA GOC ATA GAC THE GTA ACC CTG ATC CTC CCC GAC CTC TTG 122.2  2426 COC GAT ACC CTG GAT GET GTA GOC TA ACC CTG GAA GAT ACC CTG ATC CTC CTC 122.2  2426 COC GAT ACC CTG ATG CCT CAA GAT GAC GAG AGA ACC TA ACC CTG CAC ACC CTC TTG 122.2  2426 COC GAT ACC CTG CAT GC CTG AAA CTT CTG CAC GAG ACC TA CTG CTC ACC CCC GAC AAA CTG CAA CTG CTG ACC CCC GAC AAA CTG CAC ACC CTG ACC CCC GAC AAA CTG CAC CAC CAC AAA CTG CAC CAC AAA CTG CAC ACC CTG ACC CCC GAC AAA CTG CAC CAC CAC CAC CAC CAC CAC CAC CAC CA	- 4	1.	•	•	,	×	L	c	•	I	V	3	4	•							M	• • • • •
2296 AGTTANATIGCTAACGCACTCAGCACTCITICT ATO ANA TOT AND ANT GOOD CTC ATC CTC ACC 2216.2  23164 ACC STC ACC CTG GAT GET GTA GOC ATA GAC THE GTA ACC CTG ATC CTC CCC GAC CTC TTG 122.2  2426 COC GAT ACC CTG GAT GET GTA GOC TA ACC CTG GAA GAT ACC CTG ATC CTC CTC 122.2  2426 COC GAT ACC CTG ATG CCT CAA GAT GAC GAG AGA ACC TA ACC CTG CAC ACC CTC TTG 122.2  2426 COC GAT ACC CTG CAT GC CTG AAA CTT CTG CAC GAG ACC TA CTG CTC ACC CCC GAC AAA CTG CAA CTG CTG ACC CCC GAC AAA CTG CAC ACC CTG ACC CCC GAC AAA CTG CAC CAC CAC AAA CTG CAC CAC AAA CTG CAC ACC CTG ACC CCC GAC AAA CTG CAC CAC CAC CAC CAC CAC CAC CAC CAC CA	2229	ACC	ATG	144	100	141	101	CAT	c <del></del>	tca	caa	CTT A	TCAT	<b>-</b> 017	BACE			. =				
2364 ACC GTC ACC CTG GAT GCT GTA GCC ATA GC CTG ATA GCC TTG ATA GCC GTA ACC ATA GCC CTG GAC GCC ATA ACC GCC ACC ACC ACC GCC ACC ACC ACC GCC ACC A	2	Ť	ĸ	I	7	4	5	2	y	•		,,,,,,,				••••	N.	R	, X.		AICAC	1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2364 ACC GTC ACC CTG GAT GCT GTA GCC ATA GC CTG ATA GCC TTG ATA GCC GTA ACC ATA GCC CTG GAC GCC ATA ACC GCC ACC ACC ACC GCC ACC ACC ACC GCC ACC A	2296	MCT	***	TTT	7334	~~~	<b>-</b>	~					• • •		_							
2344 ACC CTC ACC CTG GAT GET GTA GGC ATA GGC TTG GTA GGC GTG GTG ACC GTG GTG GAT ACC GGC GTC GAA GTG GAT GTG GAC GGC GGC GTG GTG GAT ACC GGC GGC GGC GGC GGC GGC GGC GGC GGC	:			•		~~		-	maka Ffa	Y AT	· L	± τς 2	TAA	5 XX 1	T OC	e cr			c Ÿ1		د نحد	5101
2414 CDC CALT AFE COC CTG ATG CCT GAA CGT GAA CGT GAA CGT AAC CTG CAA CASG FET CTG CTG 2403  2180 TTU USU TGG GCA TGC CAG CAC AAC TTC TGG TCC GGT AAC CTG CTG AGC CGG GCG AAC TTC 72  2181 TTU USU TGG GCA TGC CAG CAC AAC TTC TGG TCC GGT AAC GTG CTG AGC CGG GCG AAC TTC 72  2184 ACT CCC CAT ULU UCT GTT CAC AAT TAA TCATCCGCTCGTATA ATG TGT CAA ATT GTG AGC CAG 7406  2187 TAA CCAATTCAACACACACACACACACACACACACACA	2264											_	_	_	•	_	-	•	•	_	6	- 4
2414 CDC CAT ARE CDC CTG ARG CCT GAA CGT GAC CGT AND CGT AND CGT AND CGT AND CTG CTG CTG CTG CTG CTG CTG AND CGT AND CGT CTG CTG CTG CTG CTG CTG CTG CTG CTG	13	ACC 7	erc v	ACC	CITE	GAT	्रह्म	GTA	, ecc	Ϋ́ΣΥ	COC		قمتا	Y3C	œ	GTA		ccc	aac	ctc	πc	2423
2486 TTU USU TOS GCA TOC CAO CAO CAO CAO TTO TOS TEC COST AND CTG CTG AGC COS GCA AND CTA TOT TOS TEC COST AND CTG CTG AGC COS GCA AND CTA TOT TOS TEC COST AND CTG CTG AGC COS GCA AND CTA TOT TOS TEC COST AND CTG CTG AGC COS GCA AND CTA ATC TOS TAN CANATTENANCAMENTAL AND CTG CTG ATC CAO ATT GTG AGC GCA TANK TOS TAN CANATTENANCAMENTAL AND CTG CTG ATC CAA ATT ATG AAA AAA AAC ATC AAA AAA AAA AAC CAG GTA ATC TOS		•	•		-	_			-	-	•	_	•	_	-		-	-	-	L	:	32
2486 TTU USU TOS GCA TOC CAO CAO CAO CAO TTO TOS TEC COST AND CTG CTG AGC COS GCA AND CTA TOT TOS TEC COST AND CTG CTG AGC COS GCA AND CTA TOT TOS TEC COST AND CTG CTG AGC COS GCA AND CTA TOT TOS TEC COST AND CTG CTG AGC COS GCA AND CTA ATC TOS TAN CANATTENANCAMENTAL AND CTG CTG ATC CAO ATT GTG AGC GCA TANK TOS TAN CANATTENANCAMENTAL AND CTG CTG ATC CAA ATT ATG AAA AAA AAC ATC AAA AAA AAA AAC CAG GTA ATC TOS	2424	cac	Ġ	YEC	œ	CTG	YIC	्दा	<b>CN</b>	्र व्या	CAC	COA	ध्व	MC	CMC	رعد	cac.	APG	fer	cre	CTC	2433
2944 ACT CCC CAT LLU UCT GTT GAC AAT TAA TCATCCCCTCGTATA ATC TOT GAA ATT GTG ACC GGA ACC GGA ATT GTG ACC GGA ACC ACC	13	x	0	I	R.	· <u>*</u>	R	R	Z	R	D	C	A		1					v	1	
2944 ACT CCC CAT LLU UCT GTT GAC AAT TAA TCATCCCCTCGTATA ATC TOT GAA ATT GTG ACC GGA ACC GGA ATT GTG ACC GGA ACC ACC	2184	TTC	Œ.	166	Œλ	700	CNG	CAE	w	110	700	700	507	110	~	~~~	100					
2514 ACT COC CAT CUU UCT GIT CAC AAT TAA TCATCGOCTCGTATA ATC TOT GCA ATT GTC AGC GEA 7 ATC TO THE ACC GEA ATT GTC AGC GEA 7 ATC TOT AGC AAT TOT AAA AAC ATC AAA AAA AAC CAG GTA ATC CAG CAG CAG ATC CAG CAG CAG ATC CAG CAG CAG CAG CAG CAG CAG CAG CAG CA	53	7	R		A	C	9	0	2	r		3	•	ĩ	V	L		1		I M.	_	
2607 TAA CANTTENERCHEGALIACASGATCE ATE CAT ATG ANA AND AND AND AND CAG GTA ATG 26'0  2671 AND CHO COT COC AND TOT AND CTO CTO AND AND THE AND THE AND CAG CTG ATC CAD CTG AND 2731  2731 APC CAD CTG COC AND CAC CCT ATC CCT ATC CTO ATC CTO AND CAD CTG TAT CAD CTG ATC CAD CTG TAT CAD CTG ATC CAD CAD CAD THE AND CAC CCT ATC CTO ATC CTO ATC CTO ATC CAD CAC THE CAD CAC THE CAD CAC THE CTG TAT CAD CTG THE CAD CAC THE CAC THE CAD CAC THE CAC THE CAD CAC THE CAC THE CAC THE CAD CAC THE CAC T	2 54 4	_	<b>~~</b>	-								<b></b> -				_	-	•		-		, 4
2607 IAA CHAPTTENENCHOMMANACHSCRICE RIU CAT ATO MAN AND ATC AND CAM CAG GTA ATG 26'0  2671 AMC CTC CCT CCC AND TCT ANA CTO CTC AND CAM GAM TAG AND TOC CAM CTG ATC CAM CTG AND 2131  2731 AMC CAM CTG CCT AND CTC AND CTC CCC ATC CTC ATC CTC CCC CAG CTG ATC CAM CTG AND 2132  2731 AMC CAM CTG CCC AND CTC TATC CCC CTC ATC CTC CCC CCC CAG CTG TAC ATC CCT TUT CCT 2700  2831 T R O P R A G I C L I L R R R A A I L R R R A I L R R R R A I L R R R R R A I L R R R R R R R R R R R R R R R R R R	73	~:	-	~	7	7	A	טעב	MI	IM	TCA.	1000	CTCS	TATA	716							
2871 AMC CTG CAT CGC AMC TGT AMA CTG CTG AMA GTG CTG ATC CTG AMA GAA TGC CMG CTG ATC CAA CTG AMC 2731:  2731 ATC CMA CAG TTG CAG AMA CTG CTG ATC CTG ATC CTG ATC CAG CAG CAG CAC CAG CAG CAG CAG CAG CA							6	بعق	<u>v _</u>			_		•	-	-	-	•		•	•	<u> </u>
2671 AMC CTC CCT CCC AMC TCT AMA CTC CTC CTC AMA CAA TMC AAA TCC CMC CTG ATC CAA CTG AMC 27330 13 7 8 0 9 8 8 8 2 0 2 7 8 8 9 8 8 0 2 7 8 8 9 2 8 8 2 7 8 8 9 8 8 9 2 8 7 8 8 9 8 8 8 9 2 7 8 8 9 8 8 9 8 8 8 9 8 8 8 9 8 8 9 8 8 8 9 8 8 9 8 8 8 8 9 8 8 8 9 8 8 8 9 8 8 8 9 8	2607	IM	CAA	1276	NENCE	NOON	MEM	jey i											NC G	IA A	77	2610
2731 AFC GAA CHG TTC GAA GCA CCT ATC GCT CTC ATC CTC GCC GCC CAAA AAA AAA GCA TAC GCT ATC GCT ATC GCT ATC GCT CTC ATC GCT ATC ATC GCT ATC ATC GCT ATC GCT ATC ATC ATC GCT ATC ATC ATC ATC ATC ATC ATC ATC ATC A										_		_	_	_		_	-	_	•	M		12
2731 AFC GAA CHG TTC GAA GCA CCT ATC GCT CTC ATC CTC GCC GCC CAAA AAA AAA GCA TAC GCT ATC GCT ATC GCT ATC GCT CTC ATC GCT ATC ATC GCT ATC ATC GCT ATC GCT ATC ATC ATC GCT ATC ATC ATC ATC ATC ATC ATC ATC ATC A	2671	116	CIC	<b>41</b>	cce	AAC	363	**	4	CTO	*	EAA	TAG	*	TCC	CAG	CTG	ATC	CAA	CTE	MC	273:
2791 GAT GAA GGT AAA ACC THC TGT ATG CAG TTC CAG TCC CAG ACC AAA AAA GAA CGT GTT AAC 2890 S S S S S S S S S S S S S S S S S S S	' '	W	L	6	,		•	K	L	L	X	8	Y					L		L		
2791 GAT GAA GGT AAA ACC THC TGT ATG CAG TTC CAG TCC CAG ACC AAA AAA GAA CGT GTT AAC 2890 S S S S S S S S S S S S S S S S S S S	2731	ATC	au	cae	777	GAA	ŒA	<b>41</b>	ATC	cæ	czc	ATC	<b>C70</b>	-	GAT	**	***	100	469		~	
2791 GAT GAN GGT ANA ACC THE TITE ATC CAS TTO CAS TO ANA AND ANA GGA THE ATC GAC CAC 2850  2851 GTA TOT CTG CTG TAC GAT CAS THE CTG TTG TTC CGC GGC CAC ANA ANA GAN COT GTT AND 2510  2911 CAC CTG GGT AND CTG GTA ATC ACC TGG GGC GCC CAG ACT TTC ANA CAN GAN COT GTT AND 2510  2971 ANA CTG GGT AND CTG TAC ATC GTT AND AND AND AND ACC ATC CGG AND AND CTG TTC AND 1122  2971 ANA CTG GGT AND CTG TAC ATC GTT AND AND AND AND ACC ATC CGG AND AND CTG GTT CAD 1122  2971 ANA CTG GGT AND CTG TAC ATC GTT AND AND AND AND ACC ATC CGG AND AND CTG GTT CAD 1122  2971 ANA CTG GGT AND CTG TAC CTG TAC ATC GTT AND AND ACC ATC CGG AND AND CTG GTT CAD 1122  2971 ANA CTG GGT AND CTG TAC ATC GTT AND AND AND ACC ATC CGG AND AND CTG GTT CAD 1122  2971 AND TAC TAC CTG AND CTG TTC ATC GTT AND AND ACC ATC CAT CAT CTG ATC CTG AND AND CTG TAC CTG AND TOC CAC TTC ACC TTC CAD 1122  2971 AND TAC TAC CTG AND AND AND AND AND AND TCG ATC GTA AND AND TCC CAG TCT TTC ACC TTC CAD 1122  2971 AND	33	1	R	9		2	A				L		L	•	_	<u>~</u>		ī	H COI	1		
2851 GTA TOT CTG CTG TAC GAT CAG TGG GTA CTG GTA CTG GGC GGC CAG AAA AAA GAA CGA GTT AAC 2910 73 V C L L Y D Q W V I. S B P B R B R B R A Y H D N 72  2851 CAC CTG CGT AAC CTG GTA ATC ACC TGG CGC GGC CAG ACT TTC AAA CAA CAA CGA GCT TTC AAC 3970 93 8 L G G W L V I T W G A Q T P R R Q A P P B 112  2871 AAA CTG GGT AAC CTG TTC ATC GTT AAC AAC AAA AAA AAA ACC ATC CGG AAC AAC CTG GTT CAA 1030 113 8 L A W L P I V W H W R R R T I P W W I. V P 132  3031 AAC TBW UTG ACC CGG ATG TCT CTG GCA AAC AAA AAA ACC ATC CGG AAC AAC CTG GTT CAA 1030 1133 8 T L T P B R S L A Y W P P N D D G G T AAA TGG CAC 3090 133 8 T L T P B R S L V L W T T T P R 172  3091 TAC AMG AAA AAC TCT ACC CGG ATC CTG GCA ACC CTG AAC ACC CAG TCT TTC ACT TTC CAA 1130 1133 8 V B Y L V R G L R B E Y Q L W C Y V R 192  3181 CAA GTA GAA AAC AAA CGG ATC ATC CCC CAG ATC ATC TCT CAA CTG AAC TCT TAC GTA AAA 1210 1173 8 V B Y L V R G L R B E Y Q L W C Y V R 192  3181 CAA GTA GAA AAC AAA CGG ATC ATC CCC CAG ANG ATC TTC ATG TCT LAL UTG ATC TTC TAC AAC 1270 1193 I W R B E P I I Y I D S W S Y L I Y R B 211  3271 CTG AEC AAA CGG TAC CTC ATC CCC CAG ANG ATG TAC GTG AAC CTG CAAC TCT TAC ATC TCT TAC AAC 1270 1311 CAA ACT TTC CTG AAA TA TAC ATC CCC CAG ANG ATG TAC TTC CAA CTG AAC TCT TAC TCT TAC AAC 1270 1371 CTG AEC AAA CGG TAC CTC ATC CCC CAG ANG ATG TAC TCT CAAC CTG CAAC TCT TAC TCT TAC TAC TCT TAC TCT TAC TAC	2291	<b>a</b> •	GAA	~	***	100	-	***	100											_		
2851 GTA TOT CTG CTG TAC GAT CAG TGG GTA GTA GTA CTG TCC CGG GCC CAG AAA AAA GAA CGT GTT AAC 2910 73 V C L L Y D Q W V 1. 8 9 8 8 8 8 8 7 V W 22 2851 CAC CTG GGT AAC CTG GTA ATC ACC TGG GGC GCC CAG ACT TTC AAA CAA CTA CCT TTC AAC 2970 93 8 L G G W L V I T W G A Q Y P R R Q A P R R 2911 2971 AAA CTG GGT AAC CTG TGC ATC GTT AAC AAC AAA AAA ACC ATC CGG AAC AAC CTG GTT CAA 1020 1131 8 L A W L P L V W R W E R T I P W W I I I P R I I I P R I I I P R I I I P R I I I P R I I I I	33	5	E .	ξ.	~~	1	Y	c .	M	0	P		705								-	
2911 CAC CTG CCT AAC CTG GTA ATC ACC TCG GCC CCC CAG ACT TTC AAA CAC CCT CAA CCT TTC TTC					_		-	_		-	Ĭ.	-	•		_	•	-	_		_		/2
2911 CMC CTG GGT AMC CTG GTA ATC ACC TGG GGC GCC CMG ACT TTC AAA CTG GGA ACC TTC AMC 1020 93 8	2851	GIY.	खा	èse.	CIE	iw	<b>d</b> i	Č	778	(4.2)	-	uce						an	cet	OTT	MC	2910
2971 AAA CTG GCT AAC CTG TTC ATC GTT AAC AAC AAA AAA ACC ATC CCG AAC AAC CTG GTG GAA 1020 113 E L A B L P L V E B E E T L P E H I. V P 132   3031 AMC TRU CTG ACC CCG ATG TCT CTG GCA 2AC TGC TTC ATG GAT GAT GAT GGT GCT AAA TGG GAT 1040 1133 B T L T P E S L A Y E P H D D D G G T AAA TGG GAT AAA TGG GAT 1320 1133 B T L T P E S L A Y E P H D D D G G T AAA TGG GAT AAA TGG GAT 1320 1133 B T D T P E S L A Y E P H D D D G G T TC ACT TTC ACT TTC GAA 1150 1133 B T P E B S T E B S L T P E B 172 173 B V B T L V E G L R H E P Q L B C Y V E 192 173 E V B Y L V E G L R H E P Q L B C Y V E 192 173 E P E R E P L T T T T T T T T T T T T T T T T T T	'*	•	•	•	٠.	1	U	<b>Q</b>	-	٧	1.	*	•	7		×	E	8	R	٧	•	92
2971 AAA CTG GCT AAC CTG TTC ATC GTT AAC AAC AAA AAA ACC ATC CCG AAC AAC CTG GTG GAA 1020 113 E L A B L P L V E B E E T L P E H I. V P 132   3031 AMC TRU CTG ACC CCG ATG TCT CTG GCA 2AC TGC TTC ATG GAT GAT GAT GGT GCT AAA TGG GAT 1040 1133 B T L T P E S L A Y E P H D D D G G T AAA TGG GAT AAA TGG GAT 1320 1133 B T L T P E S L A Y E P H D D D G G T AAA TGG GAT AAA TGG GAT 1320 1133 B T D T P E S L A Y E P H D D D G G T TC ACT TTC ACT TTC GAA 1150 1133 B T P E B S T E B S L T P E B 172 173 B V B T L V E G L R H E P Q L B C Y V E 192 173 E V B Y L V E G L R H E P Q L B C Y V E 192 173 E P E R E P L T T T T T T T T T T T T T T T T T T	2911	Ç	C10	व्य	MC	<b>C10</b>	ŒÀ	ATC	ACC	rœ	cec	ecc	26	ACT	TTC:	144	cac.	CAA	œ	110	we	2970
2031 AMC IMU CTE ACC CCE ATE TCT CTG GCA IAC TOG TTC ATG GAI CAR GGT GGT AAA TGG CAR 1690 1592 1593 W T L T P H S L A Y W P P H D D G G R W D 1592 1593 W T P P H S L A Y W P P H D D G G R W D 1592 1593 W T P P H S L A Y W P P H D D G G R W D 1592 1593 W T P P H S L A Y W P P H D D G G R W D 1592 1593 W T P P H S L A Y W P P H D D G G R W D 1592 1593 W T P P H S 172 1593 W T P P P H S 172 1593 W T P P P T P P H S 172 1593 W T P P P H S 172 1593 W T P P P H S 172 1593	, ,,	•	÷	G	1	L	¥	I	Ī	W	4		9								_	
2031 AMC IMU CTE ACC CCE ATE TCT CTG GCA IAC TOG TTC ATG GAI CAR GGT GGT AAA TGG CAR 1690 1592 1593 W T L T P H S L A Y W P P H D D G G R W D 1592 1593 W T P P H S L A Y W P P H D D G G R W D 1592 1593 W T P P H S L A Y W P P H D D G G R W D 1592 1593 W T P P H S L A Y W P P H D D G G R W D 1592 1593 W T P P H S L A Y W P P H D D G G R W D 1592 1593 W T P P H S 172 1593 W T P P P H S 172 1593 W T P P P T P P H S 172 1593 W T P P P H S 172 1593 W T P P P H S 172 1593	2971	ш	CZG	<b>627</b>	N.C	CTE	TE	ATC	GTT	**	110	111	222	200	١		110	110				1010
3031 AMC THE CTE ACC CCE ATC TCT CTG GCA 2AC TOG TTC ATG GAT GAT GGT GGT AAA TOG GAT 1630  3091 TMC AMC AMA AMA CTT ACC AMC AMA TCG ATC GTA CTG AMC ACC CAM TCT TTC ACT TTC CAA 3150  3181 GAA GTA GMA TAC CTG GTT AMA GGT CTB GCT AML AMA TTC CMA CTG AMC TGT TAC GTA AMA 3210  2181 GAA GTA GMA TAC CTG GTT AMA GGT CTB GCT AML AMA TTC CMA CTG AMC TGT TAC GTA AMA 3210  2773 E V B Y L V E G L R B E F Q L B C Y V E 192  2311 ATC AMC AMA AMA AMA GGB ATC ATC 2AC ATC GAT TCT ATG TCT TAL UTE ATC TTC TAC AMC 3210  183 I B R B E P I L Y L O S B S Y L F T B 312  3271 CTG ASC AMA CCC TAC CTC ASC CCC CAM AMP ATG TAC AMA CTG CCC AMC ACT ATC TCL TCL TCL 3310  2131 GMA ACT TTC CTG AMA TAA TAA TAA TAAGTCCAACCTTCCCAACCTTCCCCACCCCCCCCCC	113	L	L	À		L	ř	I	Å.			I									_	
3091 THE MAS AND AND THE MAE AND AND TOG ATC GTA CTG AND ACC CAS TOT TTC ACT TTC CAD 3150 L53 Y / B   B   S   Z   B   E   S   I   V   L   B   T   Q   S   F   T   F   B   172    3181 CAD STA STA STA COS OFF AND STA CTG AND AND TTC CAD CTG AND TOT TAC GTA AND 3210    23181 ATC AND AND AND AND COS ATC ATC TAC ATC SAS TOT ATG TOT TALL UTG ATC TTC TAC AND 3210    2311 ATC AND AND AND AND COS ATC ATC TAC ATC SAS TOT ATG TOT TALL UTG ATC TTC TAC AND 3210    2311 CTG ASC AND COS THE CTC ASC CCC CAS ASS ASS ATS THE AND CTS CCC AND ACT ATC TOLL TALL TALL TALL TALL TALL TALL TALL	1,031		-			~~							_	-	_		•					
3091 The MS AAA AAC TUT ME AAC AAA TOG ATC GTA CTG AAC ACC CAG TOT TTC ACT TTC CAA 3150 153 Y W W W W W W W W W W W W W W W W W W	133	~	I	L	1	r	2	S	L	7	Y									<u>100</u>		1
2181 CAA SEA CAA PAC CTO SET AAS SET CTS CAT AND AAA TTC CAA CTG AAC TGT TAC GTA AAA 1210 173 E V E Y C V E 192 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2		_	_	_	-	-		•	_		_	•	-	•-	_	_	•	-	-	-	•	-3/
2181 CAA STA CAA SAA COE STE AAS ST CTS COT AND AAA THE CAA CTG AAC TOT TAC GTA AAA 12210 173 8 V 8 Y L V E G L R S E Y Q L W C Y V E 192 2211 ATC AAC AAA AAA COS ATC ATC TAC SAC SAC TOT TAL UTG ATC TTC TAC AAC 1270 183 1 W R S E P I I Y I O S N S Y L I Y I S 212 2271 CTG ASC AAA CCG TAC CTC CAG ASG ATG ATG TAC CCG CAG ASG ATG ATG TAC CCG AAC ACT ATC TCU TCU TCU TCU TCU TCU TCU TCU TCU T	3091		<b>M</b>	w	WC	T.T	ME					GIA										1 4 1
2311 ATC AMC MAA AMA COB ATC ATC EMC ASC SAS TCT ATG TCT TML UTG ATC TTC TAC AMC 1270 183 I W R W W P P I I Y I D S W S Y L I Y I B 212  3271 CTG ASC MAA CCG TMC CTC ASC CCC CAS ASS ATG TAC TAC CCC CAS ASS ATG TAC TAC CCC AMA CTS CCC AMA CT ATC TCU TCU TCU TCU TCU TCU TCU TCU TCU T		• ′	-	-	-	•	4	-	•	•	Ţ	•	•		T	9	3	7	7	7	1	172
2311 ATC AMC MAA AMA COB ATC ATC EMC ASC SAS TCT ATG TCT TML UTG ATC TTC TAC AMC 1270 183 I W R W W P P I I Y I D S W S Y L I Y I B 212  3271 CTG ASC MAA CCG TMC CTC ASC CCC CAS ASS ATG TAC TAC CCC CAS ASS ATG TAC TAC CCC AMA CTS CCC AMA CT ATC TCU TCU TCU TCU TCU TCU TCU TCU TCU T	3181	CAA	<b>72</b>	<b>~</b>	TAC			144	007	CDS	œı	ANC	*	170	CNA	CTG	w	W	IAC	GZA	AAA	3213
3271 CTG ASC AAA CCG TAC CTC ASC CCC CAS ASS ATS TAC AAA CTS CCC AAC ACT ATU TUU TUU TUU TUU TUU TUU TUU TUU TU	173	8	٧	ı	¥																_	
3271 CTG ASC AAA CCG TAC CTC ASC CCC CAS ASS ATS TAC AAA CTS CCC AAC ACT ATU TUU TUU TUU TUU TUU TUU TUU TUU TU			M	**	λAC	**	<b>CD4</b>	ATC	Ate	220	AZC	411	767	ATG	TCT	TM:	(70	ATC	110	120	w	1270
3271 CTG ASC MAA CCC TAC CTC ASC CCC CAS ASC ASC TAC MAA CTG CCC AAC ACT ATU TUU TUU TUU TUU TUU TUU TUU TUU TU	2311	ATE													_		_	4	•		_	
3331 GAA ACT TTC (TO BAA TAA TAAGTCCACCTCCAGCCCACCTCCCACCCCCCCCCC	193	I		•	•	-	•															
3331 CAN NOT THE CITE AND THE TRACTICENCETECONOCCUSCOTCOCCOCCOCCOTTONOCCOCCOCCOCCOCCOCCOCCOCCOCCOCCOCCOCCOC	193	I					_															
	3271	cte Cte	) VEC	<b></b>	cœ	126	cze													_		
	3271 313	I C76 1.	ASC T	I I	ces	TAC Y	CIC	-	<u> </u>	<u> </u>	×	*	7	X	L	<u> </u>	•	1	1	8	•	
- 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1	3271 213 3331	: C76 1.	ASK T	I TTC	ces erro	TAC Y	CTC L The	TNY	, ::::::::::::::::::::::::::::::::::::	אלנאג פ	N CAN	*	7	X	L	<u> </u>	•	1	1	8	•	3404

I-SceI coding sequence of pSCM525 - Note the two amino acid. N-terminal extension as compared to genuine version of the gene.

# YARIATIONS AROUND THE 1-SCR -! SEQUENCE

		-2 H_	-1 H	1	<u>r</u>	Ŋ		5 		_ч_	1_		10	4	L	•	þ		:
K	į	1	20 K			K							30 N	1	£		:	i	À.
G	I	£	40	Ţ	<u>L</u>	<u>a</u>	2	A	I	1	<b>ર</b>	s	અ	c	E	3	<		1
c	Ħ	a	60 F	Ε	¥	ĸ	N	K	A	Y	М	0	70 H	Ψ.	ږد	L	_	٧	;
Q	٧	¥	80	3	P	P	н	ĸ	K	٤	R	¥	90 N	Ħ	L	a	N	_	¥
1	T	~	100 G	<b>A</b>	a	T	F	ĸ	н	Q	A	F	110 N	K	L	A	N		
1	٧	N	120 N	K	K	I	ı	P	N				130 E		¥	L	Ţ	2	•
3	L	A	140 Y	¥	٢	n	<u>D</u>	<u>0</u>	2	2	×	٧	150	Y	N	K	N	s	:
N	K	\$	1 <b>60</b>	٧	L	N	<b>T</b> .	Q	\$	۴	t	F	170 E	Ε	٧	Ε	Y	_	٧
ĸ	e	L	180 R	N	K	F	٥	L					190 K		N	K	N	<	,
1	1	Y	200	2	8	ä	s	Y	L	1	F	Y	210 N	L	ı	K	P	4	
ı	P	Q	220 M		Y	K	L	P	N		t			E	T	F	L	K	

Pasitions than ean he changed without affecting ensure activity (demonstrated)

positions -1 and -2 are not natural. The two amino acids are added due to cloning strategies

positions 1 to 10: can be deleted

position 36: G is tolerated

position 40: M or Y are tolerated

position 41: 8 or M are telerated

position 43: A is tolerated

position 44: Yer Mare tolerated

position 91: A is tolerated

positions 133 and 150: Lare tolerated

position 223: A and Sare tolerated

#### Changes that affect engyme activity (demonstrated)

position 19: L to &

position 38: I to S or M

position 39: G to D or R

parilion 40: L to Q

position 42: L to B

position 44: Die E. Gor H.

pesition 45: A to K or D

profition 46: Y to D

position 47: Its R or N

position 60: LmS

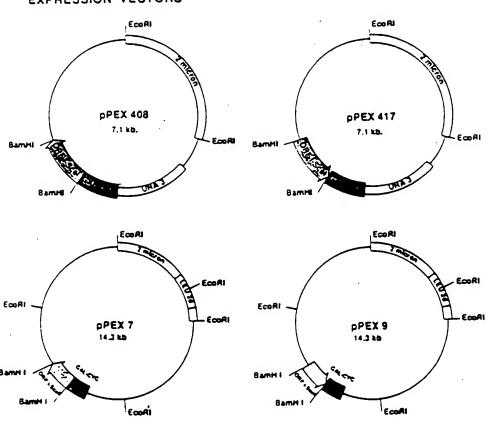
position 144:D to E position 145:D to E

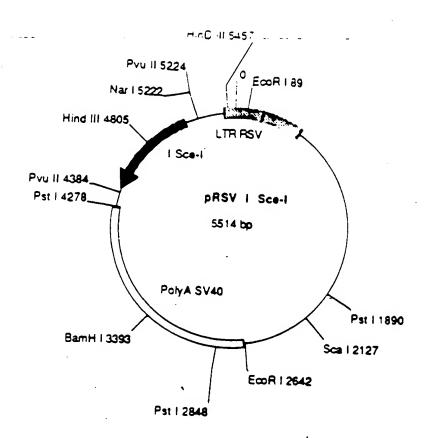
position 146 1 to E

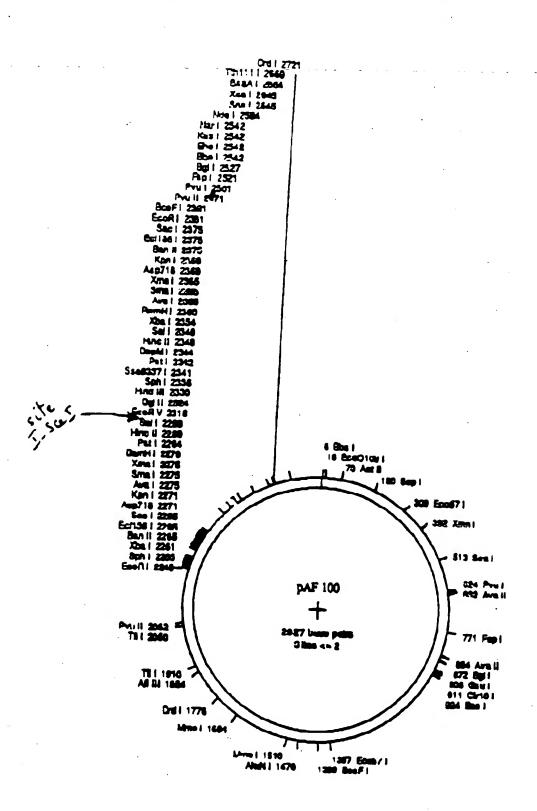
# Group I Intron Encoded Endonucleases and Related Endonucleases

	Endomuclease		
_	Extouniclease	Recognition Sequence Site Vintron	منبو
	I-6ce I (Saccharomyces mitochondria)	C G C T A G G G A T A A C A G G G T A A T A T A T A T A T A T A T	=
	I-See IV (Seecheromyces mitochondria)	SPETERT CARTAGETE TARTETALS ARGAGEASTAC	3
1	(Seccharomyces mitochondria)	CTTTGGTCATCCASAGTATACATT	•
to out	I-Cau I (Chlamydomous chlorophest)	TARCGGTCCTARGGTAGCGLAATT	A :
Llace	i-i'po i (Physarum nucleus)	A CTORGRADA GRAFTE CARES CON A A CO.	:
Two Dodecapatide Pamily for 4 hp cutteral	I-See III (Seccharomyone mitochondria) I-Ore I	CCCCAAAAUCATTUATTACTACC	: 3
	(Chlamydomonas chloropiast)	GGGTTCARANCGTCGTGAGACAGTTT	T A
Dodec	Endo See I (RF3) (Saccharomycas milochondria) (Non hormse)	CARGACATEGE TAGGETTAGTEAT	֓֞֟֝֟֟֝֟֝֟֝֟
Des	HO (Snocheromycen nucleus) Non to control	CTTTCCGCAACAGTATAATTTATAA	
	I-Cam I (Chiamydomonas mitochondria) (Ponatre endoudeus)	ACCATEEGO PCAAAT W T U T T T C T E G C C C A G T T T A C A G A A G A C C C	
	I- Pan I  (Podospacra mitochondria)  (Putative unhamateus)		
mille	(Bacceriophage T4)	-	$\dashv$
4	GAACOCICAO GITOCOAGIC	TA GAT G T T T T C T T G G G T C T A C C G T T T A A A T C T A C A A A G G A C C C A G A T G G C A A A T T	ī
	FTeV II CAAG		_
Other B	FTev III GCTATTCGTT	TITATETAPEPPPECE PEPACETTI	

# EXPRESSION VECTORS







```
paf 100 - RESTRICTION MAP
```

```
Sau 3A I
                                                                                                                                                         HOO I
                                                                                                                                          Sel I
                                                                                                                                          Mop I
                                                                                                                                           ipa ::
                                                                                                                                          V KEC
                                                                                                                                          3 stk I
                                                                                                                                        saa I
                                                                                                                                       Saa I
                                                                                                                                        ScrF 1
                                                                                                                                       ACT I
                                                                                                                                                       Boty I
                                                                                                                          Cale I TIA
                                                                                               Halk I
                                                                                                                                        Batk I
                                                                                               Ec1136 :
                                                                                                                                       BaaJ I
                                                                                                                                                                                                                                                               site
                                                                                                Bap1286 I Bon I
                                                        Sph 1
                                                                                                                                       Bon I
                                                                                                                     Ban I Bank I
                                                                                                                                                                                                Rine II
                                                                                                                                                                                           I Ega I
                                                        35p7524 I
                                                                                                                                      AVE 1
                                                                                                   ALU I
                                                        Nap I Rea I
                                                                                                                     ASP718 ALW I
           Alu I
                                                                                                                                                   11 1 1
                                                                                                                                                                                                11 1
                                                                                11 11
                                                                                                                         11 11
CCA MONTHUM ATTECOCATIONT OT AGAINST CONTROL OF THE CONTROL OF THE
COTTOCACOTTA ACCOTTA COSACA TOTOCA COCATOCOCCATOCOCCACOTOCOCA TOCOTATTO TOUGA TI AS UTOTA
                                                                           *11 31 *11 (1 11 1
                                                                                                                                                                                                11 1
            1 1 1 *
                                                         11
                                                                                                                                                                                                                                                                                                                     2318
                                                                                                                        2271
                                                                                                                                                      2279
                                                                                                                                                                                                2286
                                                         2255
             2244
                                                                                                                                                                                                        2292
                                                                                                                                         2275
                       2247
                                                                                                     2256
                                                                                                                                                  2279
                                2249
                                                                                                                                          2275
                                                         2255
                                                                                                                                                                             2284
                                                                                                  2265
                                                                                                                                                                                                 2289
                                                                                                                                            2276
                                                                                                  2245
                                                          2255
                                                                                                                                                                                                      2290
                                                                                                  2265
2260
                                                                                                                                          2275
                                                             2256
                                                                                                                                          2273
                                                                                                                                                            1280
                                                                                                  2265
                                                                                                                              2272
                                                                                                                                                          2270
                                                                                                                           2271
                                                                                                                           2271
                                                                                                                                                         2279
                                                                                                                               2272
                                                                                                                                                             2280
                                                                                                                                           2275
                                                                                                                                           2275
                                                                                                                                            2275
                                                                                                                                            2275
                                                                                                                                                2276
                                                                                                                                                2276
                                                                                                                                                2276
                                                                                                                                                2274
                                                                                                                                                 2274
                                                                                                                                                                2280
                                                                                                                                                                 2280
                                                                                                                                                                2280
```

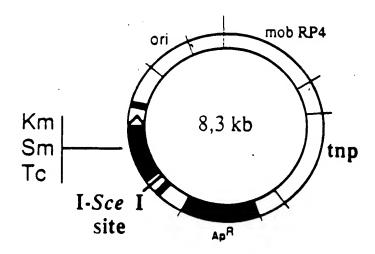
# paf 100 - RESTRICTION MAP

```
SCF I
                                                                                                                                                                                                               Acr :
                                                                                                                                                                                                               Q.F.
                                                                                                                                                                                                               374
                                                                                                                                                                                                              2 mg V
                                                                                                                                                                                                              SALK I
                                                                                                                                                                                                              2C1 :
                                                                                                                                                                                                        SCEP I
                                                                                                                                                                                                        DIA V
                                                                                                                                                                                                        RATE T
                                                                                                                                                                                                        BaaJ ?
                                                                                                                                                                                                       Ben !
                                                                                                                                                                                      SAUJA I
                                                                                                                                                                                                                                                   Sac 1
                                                                                                                                                                                    Mbo I
                                                                                                                                                                                                                                                   EgiA I
                                             Alu I
                                                                                                                                                                                    Don II
                                                                                                                                                                                                                                                   reille :
                                                                                              sfe :
                                                                                                                                       Ple I
                                                                                                                                                                                    Spa I
                                                                                                                                                                                                                            Rea I
                                Dde I
                                                                                              P#2 :
                                                                                                                                       Binf :
                                                                                                                                                                                NLA IV
                                                                                                                                                                                                                            Capé I
                   1 AUGS
                                                                                           See 337 1
                                                                                                                                                                                                                       AIT, LA
                                                                                                                                                                                BatY I
                  MD0 I
                                                                       NIA III
                                                                                                                                                                                I Bose
                                                                                                                                                                                                                       Kpa
                  DDE IT
                                                                   Spn I
                                                                                                                                                                                                                                                                                                                          Sce F
                                                                                                                                                                                Alw I
                                                                                                                                                                                                                       Ban I Alu I
                  Opa I
                                                                                                                                                                                                                                                                                                                      H 20
                                                                  BEPC 1
                                                                                                                         mine :I
                                                                                                                                                                      MAL I AVE I
                                                                                                                                                                                                                                                 9sp1286 [
              Bull I
                                                                                                                                                                                                                                                                                                                SOL ::
                                                                  #ap7524 I
                                                                                                                        ACC I SEE 1
                                                                                                                                                                                              Read 1
                                                                                                                                                                                                                                                 San II
                                                                                                                                                                                                                                                                                                                244
                                                                  Nab [ From I
                                                                                                                                                    ALW I ACK
                                                                                                                                                                                                                      ASP/LE
                                                                                                                                                                                                                                                                 Tag I
                                                                                                                                                                                                                                                                                                       Bar I
              11 1 1
                                                                  11
                                                                                                                                                   11 111 111 11
A TONANTO IA MOCITTO CA L'OCCIDENDO I CENCTE L'AGRAGA TE CE COGGA TACULA ULTURA L'ITALIGATE ESTE ESTE L'AGRAGA TE 
THE TET AGATTOGALOUTACOGACOTOCKOCTORORICTOCTROCOCCCCRICOCTCGASCTTRAGIGACCIGA MALAAA
                                                                 11 -11 1
                                                                                                                    11-1
                                                                                                                                                                  3 (1 111 11
                                                                                                                                                                                                                                                11 :-1
             2324
                                                                                                                                                                                                                                                                                                      1 11
                                                                 2336
                                                                                                      2311
                                                                                                                                                   2354
                                                                                                                                                                               2261
                                                                                                                                                                                                                     2360
                                                                                                                                                                                                                                                                  2379
                                                                                                                                                                                                                                                                                                      2387
             2324
                                                                 2336
                                                                                                                      2248
                                                                                                                                                       2356
                                                                                                                                                                                              2364
                2325
                                                                                                                                                                                                                                                                                                              2389
                                                                 2336
                                                                                                                       2348
                                                                                                                                                                    2350
                                                                                                                                                                                                  23 CB
                                                                                                                                                                                                                                                2375
                                                                                                                                                                                                                                                                                                               2369
                 2329
                                                                 2336
                                                                                                                       2348
                                                                                                                                                                             2360
                                                                                                                                                                                                                     2369
                                                                                                                                                                                                                                                   2376
                                                                                                                                                                                                                                                                                                                   2330
                2325
                                                                    2337
                                                                                                                           2344
                                                                                                                                                                             2360
                                                                                                                                                                                                                     2360
                                                                                                                                                                                                                                                                          2301
                                                                                                                                                                                                                                                                                                                       2391
                2325
                                                                                       2341
                                                                                                                                                                             2160
                                                                                                                                                                                                                     2368
                            2328
                                                                                           2347
                                                                                                                                    2351
                                                                                                                                                                             2760
                                                                                                                                                                                                                         2370
                                      2330
                                                                                           2342
                                                                                                                                    2351
                                                                                                                                                                                2361
                                          2331
                                                                                                                                                                                 2361
                                                                                                                                                                                                                                              2375
                                                                                                                                                                                 2361
                                                                                                                                                                                                                                               2375
                                                                                                                                                                                 2361
                                                                                                                                                                                                                                               2375
                                                                                                                                                                                                  2345
                                                                                                                                                                                                  2365
                                                                                                                                                                                                  2365
                                                                                                                                                                                                  2365
                                                                                                                                                                                                  2365
                                                                                                                                                                                                  236
                                                                                                                                                                                                  236
                                                                                                                                                                                                  2365
                                                                                                                                                                                                      2300
                                                                                                                                                                                                      2366
                                                                                                                                                                                                      2366
                                                                                                                                                                                                      2366
                                                                                                                                                                                                      2344
                                                                                                                                                                                                     2366
                                                                                                                                                                                                    2366
                                                                                                                           Mae III
                                                                                                  Serf I
                                                                                                  ROTE T
                                                                                                                                                                                                                                                                                                                    Pou II
                                                                                                                                                                                                                                                                                                                   Nep8 II
                                                                                           ACCORDECATE ACCIONACIONA PROSECTO CARROLA CARROLA PROSECTO POR CARROLA CARROLA
                                                                                            reconcost m i gostifanti m gostance i ce i ce i negete a me concerta a costa
             2405
                                                                                             2423
                                                                                                                                                                         2440
                                                                                                                                                                                                                         2451
                                                                                                                                                                                                                                                                                                                  2471
                                  2416
                                                                                                 2424
                                                                                                                                                                                                                         2451
                                                                                                                                                                                                                                                                                                                  2471
                                                                                                 2424
                                                                                                                                                                                                                                                   2457
                                                                                                                                                                                                                                                                                                                      2472
                                                                                                 2424
                                                                                                 2424
                                                                                                 2424
                                                                                                                           2430
```

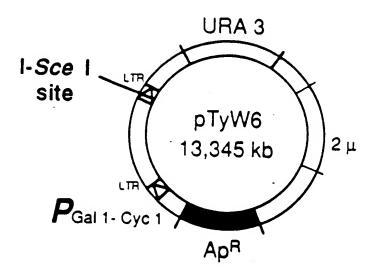
Names: pTSm  $\omega$ 

pTKm  $\omega$ 

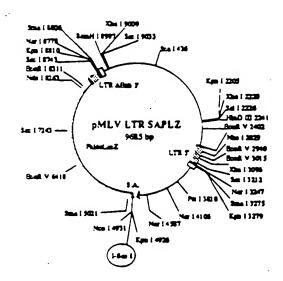
ρΤΤς ω

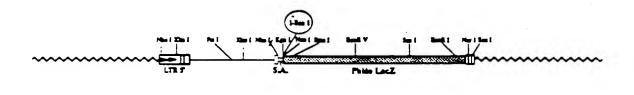


Construction: pGP 704 from De Lorenzo, with transposase gene and insertion of the linker [I-Scel] in Notl unique site

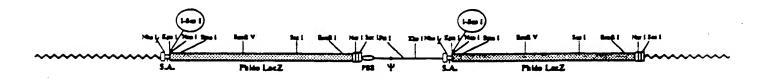


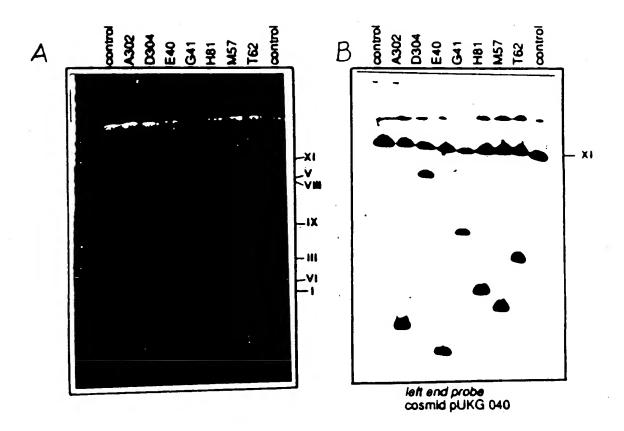
Construction: pD 123, from J. D. Boeke with insertion of a linker [I-Scel - Notl] in BamHI



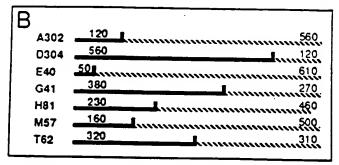


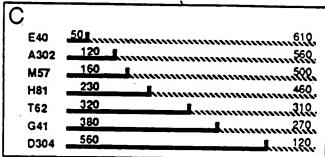


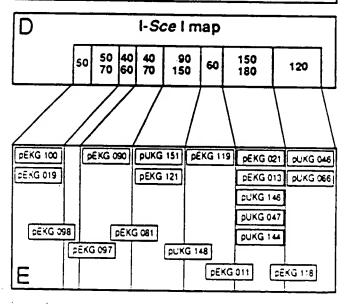


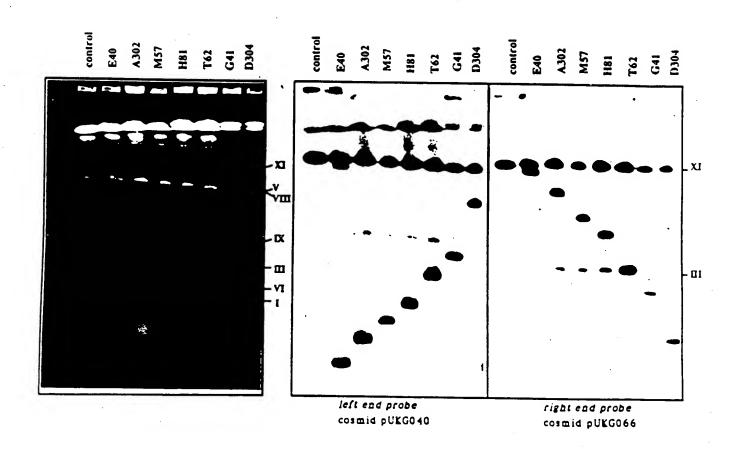


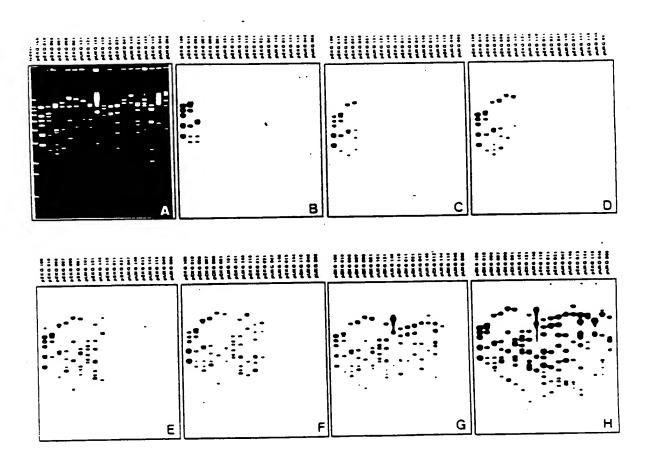
Α .	
A302	Mannen and Market Shire.
D304	
E40	.50
G41	
H81	
M57	
T62	

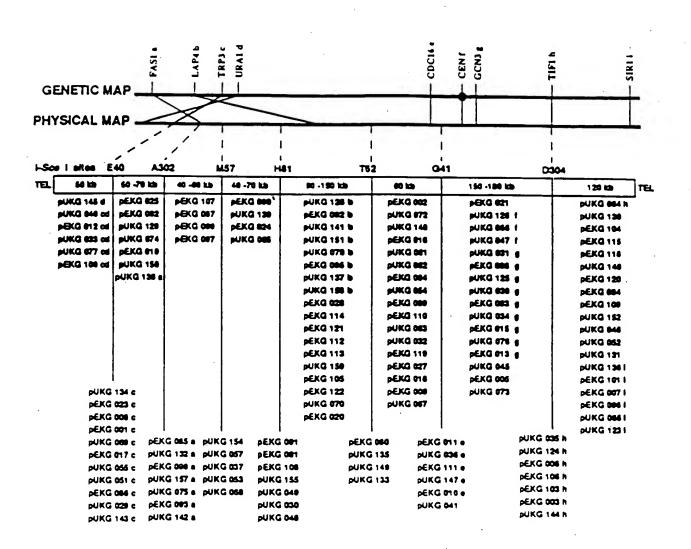


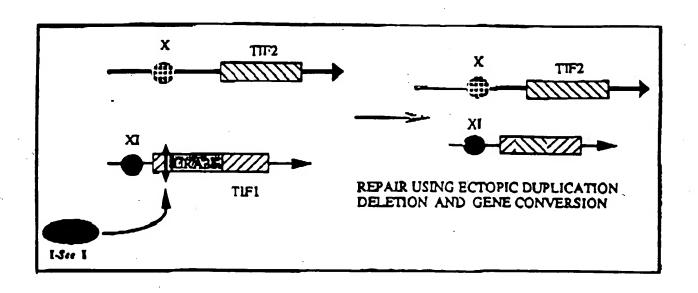


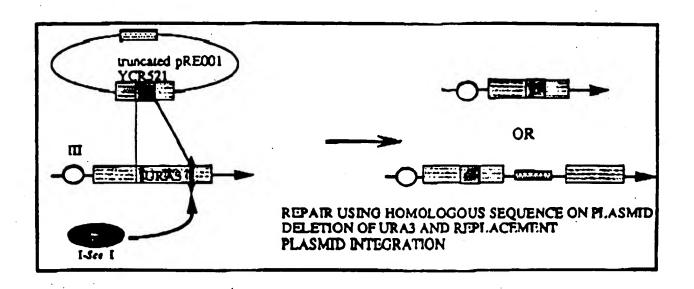


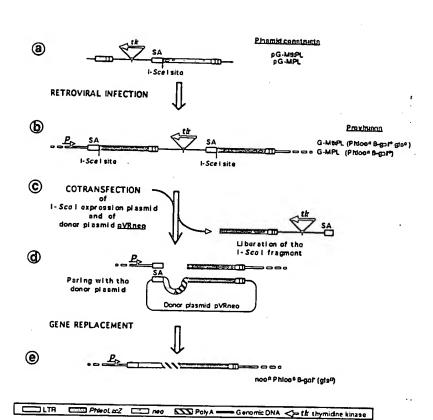




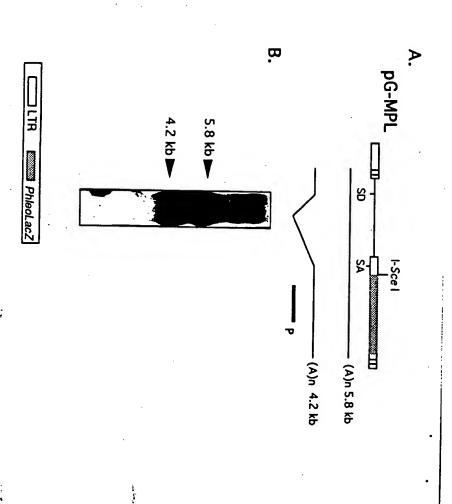


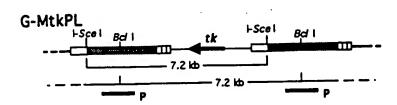






HHAN, NA

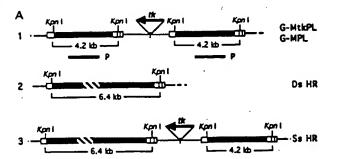


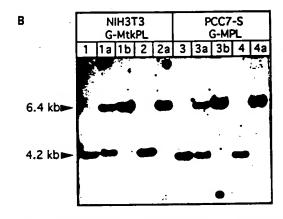


В.

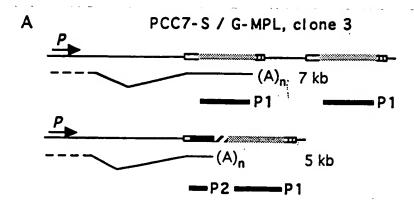
ENZYME		I <sub>2</sub> S/	e i			Ro	:/1	_
				7.6	NIIU	3T3	PCC	7-5
CELT FINE		3T3		7-5				
PROVIRUS	G-M	tkPL	_	1PL	G-M	tkPL	_	4PL
CLONE	1	2	3	4	1	2	3	4
	è	•				**	•	
7.2 kb ➤	-					•		
5.8 kb ➤			•		-	•	•	
						_		•
					·			•
							***	

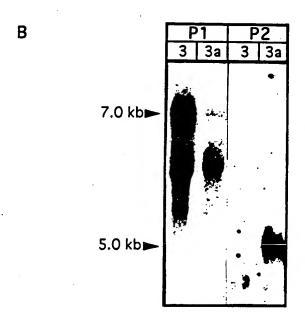
LTR PhleoLacZ tk thymidine kinase





PolyA — Genomic DNA 4-th thymidine kinase







2000

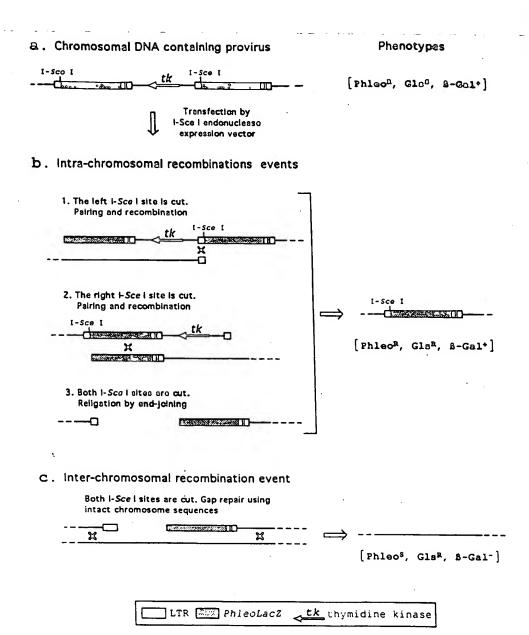


Figure 26

# a. Parental DNA, G-MtkPL



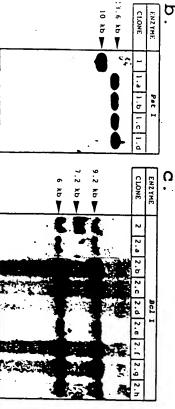
	10 kb <b>▼</b>	CLONE	ENZ YME	7,
_		1 1.1 1.2 1.3 1.4 1.5	Pat I	
7 kb	9 kb <b>Y</b>	CLONE	ENZYME	)
		2 2.1 2.2 2.3 2.4 2.5 2.6	Pat I	

# 1. Parental DNA, G-MtkPL



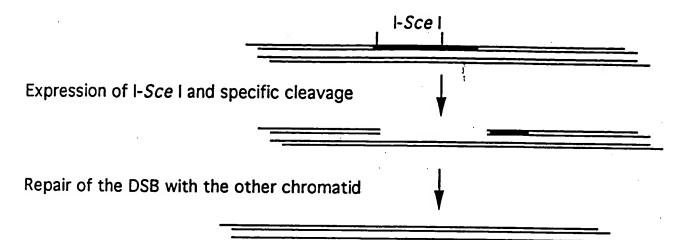
# 2. Intra-molecular recombination event





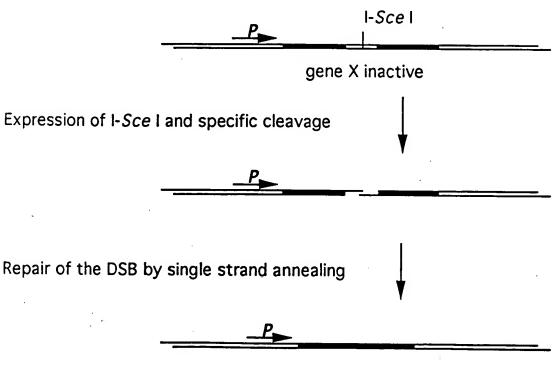
# LOSS OF HETEROZYGOSITY

Integration of artificial site or presence of specific site



# CONDITIONNAL ACTIVATION (Tandem repeat)

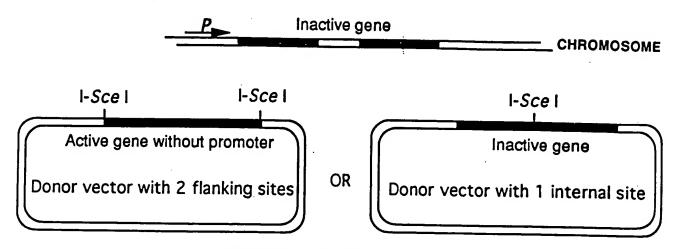
Integration of artificial site between tandem repeats



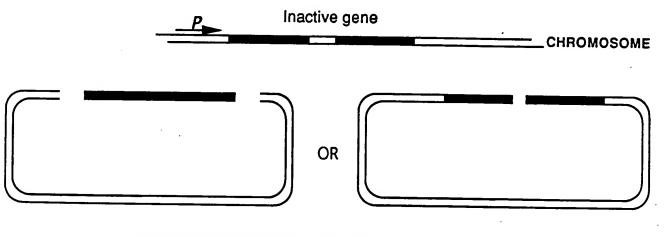
gene X active



Integration of artificial site or presence of specific site



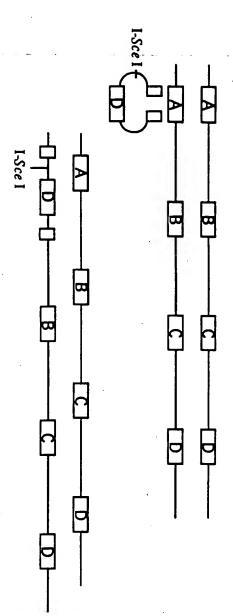
Expression of I-Sce I enzyme and specific cleavage of the donor plasmid



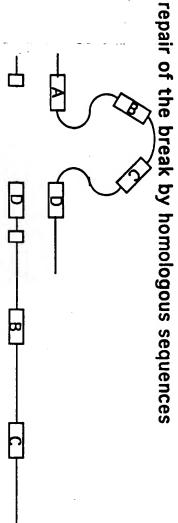
Recombination between the chromosome and plasmid



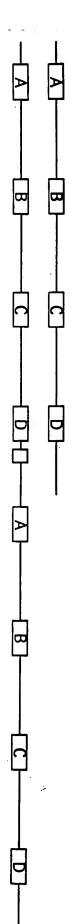
1 insertion of I-Sce I site by classical gene replacement



Specific cleavage by I-Sce I enzyme and



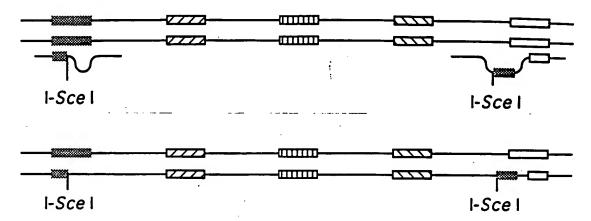
3 Duplication of the totality of the locus



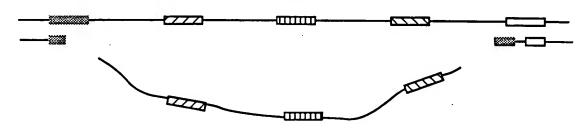




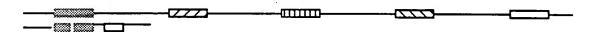
# 1 Insertion of two I-Sce I sites flanking the locus



# 2 Expression of the enzyme and cleavage



3 Recombination between the two ends



4 deletion of the locus

